



Queen's Economics Department Working Paper No. 36

IS THE COEFFICIENT OF CROSS ELASTICITY AN APPROPRIATE MEASURE FOR THE CLOSENESS OF SUBSTITUTES?

Klaus Stegemann
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-1971

IS THE COEFFICIENT OF CROSS ELASTICITY AN APPROPRIATE
MEASURE FOR THE CLOSENESS OF SUBSTITUTES?

Klaus Stegemann
Queen's University

Discussion Paper No. 36

IS THE COEFFICIENT OF CROSS ELASTICITY AN APPROPRIATE
MEASURE FOR THE CLOSENESS OF SUBSTITUTES?

Klaus Stegemann*

I. Questions to be Discussed

Authors of price theory texts usually state that the coefficient of (price) cross elasticity of demand is a measure for the "closeness" of substitutes or complements: for the closeness of substitutes when the sign of the cross elasticity between two commodities is positive and for the closeness of complements when the sign is negative. The question of sign will concern us later (Section V below). The main point I want to discuss is largely symmetrical for complements and substitutes and it is more relevant for the latter. I will therefore develop my argument in terms of substitutes. The central question is: How should we measure the relative closeness or the "degree of substitutability" of substitutes when we are dealing with more than two goods? To be sure, we always need more than two goods, if we want to measure closeness of substitutes. The statement "margarine is a close substitute of butter" makes no sense unless we have some other commodity

*The author would like to thank his colleagues at Queen's University, especially R.G. Eipsey and D.A. Vardy, for helpful criticism of two earlier versions of this paper.

or commodities in mind which are not "close" substitutes of butter, or are "less close" than margarine. As it turns out there are several alternative ways of relating two or more substitutes to a commodity. And it will be shown that, for some purposes, some of these approaches are ambiguous ways of measuring the degree of substitutability.

I should mention that, for the sake of an orderly discussion, the following section (II) starts off with an unproblematic case. The origin of this reconsideration is, of course, that I stumbled upon an inappropriate way of comparing coefficients of cross elasticity.

Section II contains the main argument of the paper as well as a few minor points which are relevant for the interpretation of coefficients of cross elasticity. In Section III, I reconcile my critique with the debate between Triffin, Chamberlin, Bishop and others on the use of cross elasticities for the classification of market forms; Section IV deals with alternatives to the concept of cross elasticity; and in Section V I relate the coefficients to the Hicksian distinction between substitutes and complements. In the final section (VI) I comment on the use of cross elasticities in Antitrust cases.

11. Should We Compare Relative Changes in Own Sales or Relative Changes in Sales of Others?

To make a simple matter easy to read and also in order to convey its practical implications, I will discuss the problem in terms of a concrete example. Let us suppose a national association of butter producers wanted to determine which of two potential substitutes, margarine and peanut butter, is the closer substitute for butter. One way of doing this is to assume that the prices of margarine and peanut butter vary within some narrow range while the price of butter and all other demand conditions remain unchanged. Using an arc formula, we would then have to compare the following cross elasticities.¹

(1)

$$\epsilon_{Q_b, P_m} = \frac{\frac{Q'_b - Q_b}{Q_b + Q'_b}}{\frac{P'_m - P_m}{P_m + P'_m}} \quad \text{and}$$

$$\epsilon_{Q_b, P_g} = \frac{\frac{Q'_b - Q_b}{Q_b + Q'_b}}{\frac{P'_g - P_g}{P_g + P'_g}}$$

1. The notation should be self-explanatory with the exception that in the second formula I use the subscript "g" (for "ground-nut butter") in order to avoid using little "p" as a subscript.

We thus compare the percent changes (shifts) in the demand for butter which result from a one-percent change in the price of margarine or the price of ground-nut butter respectively. The substitute which has the larger coefficient is the closer substitute for butter. This way of comparing coefficients of cross elasticity is unambiguous as far as it goes. Presumably, this is the approach which many textbook authors have in mind because it follows from the Marshallian way of considering "other prices" as arguments in the demand function for any one commodity. However, textbook authors never warn the reader that this is the only valid approach. They hardly ever deal with more than one "other" good at a time, for that matter.²

Our association of butter producers might be dissatisfied with the comparison employed in formulae (1). They might either not have the data to estimate the coefficients with respect to changes in the prices of margarine and peanut butter. Or they might be more eager to find out which of the two is the closer substitute when the price of butter changes and the other prices are assumed to remain unchanged. For this question, it is tempting to define the relevant cross

2. The formulae (1) approach is stated explicitly by Samuelson [13, p. 105] and Scherer [14, p. 459].

elasticities in the following way:

(2)

$$\epsilon_{Q_m, P_b} = \frac{\frac{Q'_m - Q_m}{Q_m + Q'_m}}{\frac{P'_b - P_b}{P_b + P'_b}} \quad \text{and}$$

$$\epsilon_{Q_g, P_b} = \frac{\frac{Q'_g - Q_g}{Q_g + Q'_g}}{\frac{P'_b - P_b}{P_b + P'_b}} \quad \text{respectively.}$$

This approach is particularly tempting because it runs parallel to the way in which Hicks [8, Chapter III] relates different commodities to each other. He varies the price of one good, say butter, and looks at the changes in demand for several other goods. As is well known, Hicks employs his apparatus of substitution and income effects rather than coefficients of cross elasticity and no critique of his approach is intended here. Hicks could even argue that in his framework it would not matter which way one compares the closeness of substitutes for butter, because for each substitute the absolute "cross effects" are equal under both formulae (1) and (2) as long as the conditions for his "reciprocity theorem" are fulfilled [9, Chapter XIII]. For cross elasticities we do not stipulate those conditions. What is more important, we compare relative changes in demand

rather than absolute cross effects when we employ coefficients of cross elasticity. Presumably, the elasticity approach was adopted in order to avoid a problem of units of measurement.

If we employ formulae (1), i.e. we compare relative changes in "own" sales of butter, the difference between absolute and relative quantity changes does not affect the interpretation of the coefficients as a measure for the degree of substitutability. However, a closer look at formulae (2) reveals that relative changes in sales of substitutes imply a misleading comparison: The denominators are identical, since we want to relate the effects on the demand for the two alternative substitutes to the same change in the price of butter. The numerators, i.e. the per-cent changes of quantities demanded, thus determine which of the two alternative substitutes, margarine or peanut butter, appears to be a closer substitute for butter if we judge closeness by comparing the size of the coefficients of cross elasticity. Now it is trivial to point out that the value of the numerators depends both on the absolute changes in quantities demanded and on the base quantities to which we relate the changes in demand caused by changes in the butter price. With a given absolute expansion of demand for margarine, the value of the coefficient depends on

the size of the margarine market prior to the change in the price of butter. The larger the market, the lower the coefficient for margarine. The same argument, of course, holds for peanut butter and any other good.

If the size of the market influences the value of the coefficients for individual goods it may also distort a comparison between them. We could have two historical situations where the response of consumers to a rise in the butter price is exactly the same in terms of switching to substitutes for butter and still the coefficients would give a different impression if the relative sizes of the markets for substitutes would differ in the two situations. It is also easy to construct examples in which the absolute increase in quantity (measured in tons per month) is much smaller in the peanut butter market than in the margarine market, and still the coefficient of cross elasticity makes peanut butter appear to be a much closer substitute for butter than margarine. We just have to make the peanut butter market small enough relative to the margarine market. And I do not see any good reasons why such cases are unlikely to occur in the real world.³

3. For the extreme cases, i.e. coefficients of cross elasticity equal to zero or infinity, the measure is unambiguous because the initial size of the other markets does not really matter. However, in those cases we would no longer be concerned with relative closeness of various substitutes. Substitutability would be a matter of either or, rather than a matter of degree.

It is implied in the above critique that absolute changes in the margarine and peanut butter markets would be a more relevant measure for comparing the closeness of substitutes. The measure appears to be more relevant in the sense that butter producers as a group would consider margarine as the closer substitute if, after an increase in the price of butter, they expect to lose more sales to margarine than to peanut butter no matter what the percentage changes in those markets. The butter lobby would be guided by the absolute quantities (or percentages of own sales lost) in its efforts to procure protective measures against margarine and against peanut butter. A policy maker would feel in this situation that the availability of margarine (at a constant price?) is a more important check on the price policies of a butter monopoly than the availability of peanut butter. The difficulty is, of course, that we do not always feel confident that we should use a comparison of absolute changes, because we have to measure those changes in different units. I shall return to this problem in Section IV below.

Strictly speaking, our cross elasticities are defined as partial elasticities, i.e. we assume that in the case of formulae (2), the prices of margarine and peanut butter remain unchanged when the price of butter

is changed. Difficulties with this concept arise when supply is not perfectly elastic in those markets where we measure the quantity changes. Assuming that estimation problems could be solved, the question is whether we should compare horizontal shifts of the demand curves for various substitutes or whether we should rather compare the effects on equilibrium quantities. The answer, I think, depends on the purpose for which we intend to determine the degree of substitutability. This will become clearer in later sections (III and VI). It would appear that for our association of butter producers the horizontal shift of a substitute's demand curve is of interest only if the supply of the substitute actually is perfectly elastic. If it is not, the changes in quantity bought seem to be a more relevant measure for the relative closeness of substitutes than the "pure demand" relationships. In the extreme case, it seems that butter producers need not worry about losing customers to margarine producers as long as the latter have no reserve capacities and are unable or unwilling to expand output. On second thought, this is not correct. When the demand for margarine shifts upward and the price of margarine is increased, some former buyers of margarine switch to substitutes for margarine and some former buyers of butter share the available margarine supply. Thus the

observed increase in margarine sales understates the volume of sales switched from butter to margarine whenever the margarine supply is not completely elastic. An analogous argument applies for other substitutes.

A comparison of closeness is distorted when we apply the partial elasticity formulae (2) or corresponding absolute measures in cases of non-horizontal supply. We cannot measure how much butter is substituted by the various substitutes even if some technical conversion rates allow us to express quantities of substitutes in terms of butter. If we employ formulae (1) where we measure all quantity changes in terms of the "own" commodity, butter, an upward-sloping supply curve for butter would also reduce the values of the coefficients when we interpret them as measures of the cross impact on market equilibrium. However, we could still use a comparison of such hybrid coefficients to determine the relative closeness of substitutes, because we measure all quantity changes in the butter market. Margarine, for example, could be called a closer substitute than peanut butter in the sense that a one-percent price change of margarine under the given conditions has a greater effect on butter sales than a one-percent price change of peanut butter. Such a comparison makes sense no matter what the elasticity of butter supply as long as it is not zero over the relevant range. (In any case

we have to make allowance for different time horizons.) How about cases where we cannot speak of "supply curves" because the affected firms or groups do not take their prices as given? In such cases we may have to adhere more strictly to the concept of partial cross elasticity. This will be explained after I have introduced a different purpose for the cross elasticity measure in the following section.

Before I close this section on the relative merits of formulae (1) and (2), I must point out that the formulae (1) approach shares certain shortcomings with all other cross elasticity concepts. These shortcomings in themselves severely limit the usefulness of cross elasticities as measures for the closeness of substitutes. One such shortcoming is that we only compare cross relationships for certain sets of prices, and there is no justification to extrapolate information to other price situations. One could, for instance, imagine that at a low butter price, peanut butter and butter would be closer substitutes than margarine and butter, whereas for some higher butter price peanut butter and butter are completely unrelated goods because at that price level buyers consume only peanut butter in all uses where the two products would be potentially interchangeable. Another shortcoming of cross elasticities is that they

measure the cross responsiveness of demand in terms of price responsiveness only. Again, there is no justification to extrapolate results from price to non-price parameters, unless changes in non-price parameters of a seller's offer can be translated into changes in relative prices for buyers.⁴

Finally, it must be said that cross elasticities are only unambiguous measures for the degree of substitutability to the extent that we can resort to an unambiguous rule for the definition of products. Take the formulae (1) approach: If we lump together all brands of margarine, the coefficient would be larger than for each brand taken separately. We might find that, margarine as an aggregate is a closer substitute for butter than peanut butter and yet peanut butter as an aggregate may turn out to be a closer substitute for butter than the closest brand of margarine. Now the formulae (2) approach: There is no distortion of the relative degree of substitutability measured in terms of relative changes of sales of substitutes as long as there is little variation in the coefficients for the aggregated sub-products. To the extent that there is variation, there is distortion

4. It is interesting to note that practical attempts at measuring the closeness of substitutes often emphasize non-price parameters. A United Nations study of steel substitutes [18], for example, closely examines the technical factors favouring or impeding the replacement of steel by various substitutes in various uses.

from aggregation. Of course, we found earlier that formulae (2) implies an inappropriate comparison anyway. In the context of the following section, where both formulae (1) and formulae (2) have a purpose, the problem of aggregation does not arise because products are defined as the output of individual sellers; each is assumed to be a seller of only one homogeneous product.

III. The Coefficient of Cross Elasticity as a Criterion for Isolation

The well known debate on Triffin's and Chamberlin's use of cross elasticities to classify competitive relationships between firms, a debate which was stimulated more than moderated by R.L. Bishop,⁵ centred on difficulties with the coefficient of cross elasticity that arise in the context of market form classification. I do not intend to take up that debate again. However, it is interesting to note that both our formulae (1) and (2) have been used for the classification of market forms,⁶

5. Cf. Bishop's article [2] in the American Economic Review, December 1952, and the ensuing discussion in the same journal, December 1953, June 1955, as well as Southern Economic Journal, 1959-1961.

6. Bishop [2, p. 785] "in anticipation of some later discussion of asymmetries" prefers to pitch his article primarily in terms of our formulae (2) approach. Triffin [17], in his test for "circular relationship" generally favours pairing a cross price elasticity of formulae (1) type with a cross quantity elasticity. Chamberlin [7, p. 84] introduces the formulae (1) approach and I fail to see why he did not employ the formulae (2) approach as well.

and I will explain why the difficulties which make formulae (2) inappropriate for comparing closeness do not arise in the context of market classification. The basic reason is that the named authors ask a different question.

The concept of "closeness" which has been discussed so far asks for the amounts of butter sales which are lost to or gained from the margarine and the peanut butter markets when relative prices are changed. For that concept we do not want to compare the percentages by which margarine and peanut butter sales are changing. That is why formulae (2) cross elasticities are misleading. Triffin, Chamberlin, and Bishop are concerned with a different concept of market relationships. They essentially ask the question whether or not an individual seller (or a cooperating group) has reason to expect that other sellers might react to his actions in the market. In the context of market forms, the coefficient of cross elasticity is employed as a criterion for "isolation" or "non-isolation" of an individual seller. The aim is to distinguish monopolistic and perfectly competitive relationships from oligopolistic relationships. The hypothesis is that other sellers would be expected to react to a seller's price change if their sales were affected to an "appreciable" extent and that they would be assumed to stay put if the effect on their sales were "negligible".

Provided one accepts the hypothesis and finds some operational definitions for "appreciable" and "negligible" it makes sense to measure the quantity effects in terms of percentage changes of the affected firms' sales, i.e. the formulae (2) approach is appropriate. Chamberlin [7, p. 86] sums up: "The correct requirement for isolation must remain simply that $\frac{\Delta q_i}{q_i}$ be very small for plausible variations in P_j ." Triffin [17, p. 101] and Chamberlin [7, p. 85] explicitly state that they use a relative measure to take account of possible "asymmetry" between firms as to size.

As I mentioned already, both our formulae (1) and (2) are appropriate as criteria of isolation and often both are necessary to complete a "circular test". We could use formulae (2) to check whether and by whom a butter marketing board must expect reactions to a change in the price of butter. If the percentage quantity effects on all sellers of margarine, peanut butter, and other substitutes are negligible then we can conclude that the marketing board is in an isolated position. To the extent that sellers of substitutes are noticeably affected its position is not isolated (oligopolistic relationships). We can then use formulae (1) and check whether or not the sales of the butter marketing board would be noticeably affected by price changes of

any other seller, taken one at a time. The coefficients between the marketing board and any other seller would not ordinarily be the same for both approaches, even if we assume the same relative price change. Even if the absolute quantities moving between butter and a particular substitute were the same, the coefficients under the two approaches would still be different except for the limiting case of equal market size before the price change. Thus it is quite conceivable that the formulae (1) approach leads to a different classification than formulae (2). It all depends on the number of substitutes and the relative size of competing sellers.⁷

Under the isolation test, we are interested in hypothetical changes in sales, the changes which would occur if the affected sellers held their prices constant. The rationale of the test implies that in non-isolation cases the affected sellers are not expected to stay put, and a seller contemplating a price change would definitely want to know to what extent affected firms would actually

7. If, for example, a margarine producer is very large relative to the butter board and the quantities switched do not vary too much under both formulae, we may find that under formulae (2) the margarine seller is not noticeably affected by a change in butter price, whereas under formulae (1) an equivalent change in the margarine seller's price would affect the smaller butter sales to such an extent that a high coefficient indicates an oligopolistic relationship between the two sellers.

allow their sales to change if their demand were affected by the contemplated price change. However, that obviously is a much more complicated question. The "as if" shifts in demand which determine the values of the coefficients are only one element in the affected sellers' decisions. High coefficients are assumed to be just a proxy for recognized interdependence between existing sellers.

My intention in this section was merely to distinguish the concept of isolation from the concept of closeness. Having come this far, I realize that I must at least mention one major difficulty which severely mars the seemingly straightforward classification scheme for market forms. The difficulty is that coefficients of cross elasticity do not permit the separation of "pure" competition from other cases in an unambiguous way. Triffin [17, p. 103] classified pure competition together with pure oligopoly in his general category "homogeneous competition", characterized by coefficients approaching infinity. He then [17, pp. 104-105] introduced his test for "circular relationship" to distinguish homogeneous oligopoly from homogeneous atomistic competition. Chamberlin [6, p. 265] as well as several other authors argued that under pure (atomistic) competition the coefficients of cross elasticity between individual sellers would approach zero. He suggested to employ the

straight price elasticity of demand for an additional test in order to distinguish cases of pure (atomistic) competition from other cases of isolated selling. Bishop [2, p. 797] concluded that "Chamberlin's use of E_{ji} and E_{ji} breaks down completely because large values of E_{ji} do not necessarily imply oligopolistic interdependence." He excluded the limiting case of strictly homogeneous goods from his own classification scheme, reverting to the number of sellers as criterion for pure competition [2, p. 799].

Chamberlin in 1957 [7, pp. 90-91] agreed that it seems advisable to avoid the concept of cross elasticity for the limiting case of pure competition.⁸ Moreover, in a complete turnabout, Chamberlin proposed to replace the coefficient of cross elasticity as a criterion of isolation by his measure $\frac{\Delta q_i}{q_i}$ quoted above. This turnabout was prompted by Bishop's criticism [2, pp. 789-792] of invoking supply restrictions to modify the interpretation of cross elasticities. Chamberlin agreed in 1957 that the coefficient of cross elasticity defined as a derivative and as a "pure demand" relationship was not

8. Chamberlin [7, p. 90] put the problem in the following way: "If with (given) large numbers we let demand elasticities approach ∞ , the coefficient approaches ∞ ; but if with (given) high elasticities we let the number of sellers approach ∞ , the coefficient approaches 0. The ambiguity of the result seems inherent because of the requirement for pure competition, which includes both (1) product homogeneity and (2) large numbers."

affected by his earlier stipulation of rising marginal cost or limited capacity, and he concluded that this makes the coefficient defective as a criterion of isolation. What matters is the impact on the other sellers' sales of a "feasible" price change which one seller contemplates in the light of his own demand and cost situation. "For example, it is consistent with a coefficient of ∞ that $\frac{\Delta q_j}{q_j}$ should be negligible because of the rising cost curve of firm j , which limits the magnitude of Δq_j (and hence of Δq_i) as firm j shades its price." [Chamberlin, 7, p. 86]. A seller in (almost) pure competition expects that his actions have only a negligible impact on each of many other sellers of virtually the same product, because he is unwilling to serve the whole market even if he could theoretically attract all other sellers' customers by shading his price by a very small amount. If the individual seller's profitable capacity is small in relation to the whole market for the homogeneous good, he is in an isolated position. If, on the other hand, his cost situation is such that it would be profitable for him to attract large shares of his competitors' sales, the market relationships would have to be classified as non-isolated or oligopolistic. Authors like Pfouts and Ferguson [12] seem to aim in the same direction when they distinguish "Market Classification Systems in Theory and Policy."

To my knowledge, the coefficient of cross elasticity as a means of market-form classification has never been employed for any policy purposes. The Triffin approach might have had a lasting influence, however, by corroborating the widely held view that competition is the more active (and beneficial) the closer actual market situations approach the "ideal" of homogeneous competition. The crude identification of very high Triffin coefficients with very good performance has persisted in the text books, notwithstanding the work of J.M. Clark, M.A. Adelman and others who showed that, especially in oligopoly situations, various elements of heterogeneity are essential in stimulating and maintaining active (and beneficial) competitive processes. The coefficient of cross elasticity as a measure for the relative closeness of substitutes (rather than as a criterion of isolation between pairs of sellers) was used in United States Antitrust cases on several occasions. I will return to this subject in Section VI below.

IV. Absolute Cross Elasticities and Coefficients of Mobility as Alternative Measures of Closeness

In Section II above, we found that the formulae (2) approach is misleading as a measure for the relative closeness of substitutes. If butter producers wish to determine whether margarine or peanut butter is a closer

substitute for butter when the butter price changes, there are two alternative measures which do not suffer from the shortcomings of the formulae (2) approach. Both of these alternatives have been proposed in the literature; however their specific advantages or disadvantages have not been discussed in the present context.

Boulding [3, p. 200] mentions the concept of an "absolute cross elasticity of demand" which he defines as "the absolute change in the quantity of A demanded which would result from a unit change in the price of B, other factors held constant." This definition is analogous to the Hicksian "cross effects" [9, p. 127]. In terms of our previous example, we would have to compare the following two measures of absolute cross elasticity:

(3)

$$\alpha_{Q_m, P_b} = \frac{Q'_m - Q_m}{\frac{P'_b - P_b}{P_b + P'_b}} \times 200 \quad \text{and}$$

$$\alpha_{Q_g, P_b} = \frac{Q'_g - Q_g}{\frac{P'_b - P_b}{P_b + P'_b}} \times 200$$

For reasons of symmetry with the other concepts, the denominators in formulae (3) were written as a one-percent change in the butter price rather than as a unit change. This need not be a matter of concern, however,

because we relate the quantity changes in both substitute markets to the same price change for butter. In the case of formulae (1) we must use a percent price change in order to avoid a problem of units in the denominators. In the case of formulae (3) the problem of units arises in the numerators. The two coefficients only answer the question for the relative closeness of butter substitutes unambiguously if we can compare units of margarine and peanut butter on a one-to-one basis. It seems to be accepted that the problem of units can be neglected when one is dealing with differentiated products which have a common name, e.g. different brands of coffee or brands of cigarettes or even automobiles. We almost daily use per-cent figures for market shares or percentage registration figures for such differentiated products which strictly speaking are of the "apple-and-pear" type, the idea being that consumers substitute them pound per pound or car for car. If they do not, one could try to employ conversion factors. The changes in the demand for margarine and peanut butter, for instance, could be converted into tons of "butter equivalent". This approach definitely has its limits. Had I used butter versus margarine and eggs or even butter versus margarine and electricity (as a substitute for human kinetic energy and as a fuel for heating) the problem would have been more apparent.

If we want to avoid the distortions of the formulae

(2) approach and also dislike the use of absolute quantity changes for a measure of closeness, we can obtain unit-free coefficients by relating the quantity changes to the price changing group's or firm's sales. The Bonn economist Wilhelm Krelle [10, pp. 8-12] has employed this approach for constructing his measure of "mobility of demand" (Beweglichkeit der Nachfrage).

(price) mobility of demand between butter and margarine, for example, would be defined as the percentage change in butter sales lost to (gained from) the margarine market in consequence of a one-percent rise (fall) in the price of butter. We can determine the relative closeness of the two butter substitutes, margarine and ground-nut butter, by comparing the following two coefficients of mobility of demand:⁹

(4)

$$\beta_{Q_{bm}, P_b} = - \frac{\frac{\Delta Q_{bm}}{Q_b + Q'_b}}{\frac{P'_b - P_b}{P_b + P'_b}} \quad \text{and}$$

9. Krelle [10, p. 9] points out that, alternatively, his coefficients could be defined with respect to changes in the prices of the substitute goods. This measure then is identical to our formulae (1) approach above. One could employ a combination of both concepts for a "circular test" of isolation.

$$\beta_{Q_{bg}, P_b} = - \frac{\frac{\Delta Q_{bg}}{Q_b + Q'_b}}{\frac{P'_b - P_b}{P_b + P'_b}}$$

These definitions are strictly analogous to the "own" price elasticity of demand. Under the assumption that the price changes considered have no income effects, coefficients of mobility are sectoral own elasticities, and the sum of the coefficients for all other goods (savings included) is equal to the own elasticity of butter with respect to the given change in the price of butter.¹⁰ The formulae (4) concept appears to be an unambiguous approach to measuring relative closeness of substitutes. However, several conceptional difficulties arise when we imagine how coefficients of mobility could be measured.

Firstly, there could be a problem of units in the numerators although the definition seems to avoid this problem. How would one go about measuring the shares of

10. A warning seems to be in order: Such sectoral elasticities are not the elasticities of demand in sub-markets which sellers would like to know for purposes of price differentiation. If we wanted to measure the elasticity of demand for the sub-market in which butter competes with margarine we would have to relate ΔQ_{bm} to the butter sales in that sub-market rather than to all butter sales.

butter sales being lost to (or gained from) margarine, cheese, electricity, and other products when the relative price of butter changes? Presumably, one would measure the quantity changes in substitute markets and then express them in terms of butter equivalents. Secondly, one could also encounter the problem discussed for cross elasticities in Section II that other prices might not actually remain constant when the price of butter is changed, i.e. one could not measure a pure demand relationship. Thirdly, it must be recognized that the coefficients are the smaller the narrower we define a substitute. This problem arises when we are dealing with several similar substitutes which easily can be aggregated. We would, for instance, find that the coefficients of mobility are smaller for each individual brand of margarine than for all brands of margarine taken together, although - in a food chemist's opinion - certain brands might be more similar to butter than the "average" brand of margarine.¹¹ This problem can also occur with formulae (1) and (3). Fourthly, it seems to me that Krelle is only concerned with the substitution effects of a price change and that he completely disre-

11. Krelle developed the coefficient of mobility as a criterion for isolation. In that context, a "product" is defined as the output of an individual producer. An aggregation problem does not arise as long as each firm only sells one homogeneous product.

gards potential income effects. He speaks of changes in demand "going to" or "coming from" substitute markets and seems to imply that a seller who lowers his price only gains the sales which other sellers lose. To the extent that a price cut has noticeable income effects, the losses in sales of normal substitutes will be smaller than the pure substitution effects and the gains of the price cutting seller will be larger than just the sum of sales switched from substitutes. I will have to say more on this last point in Section V below. We run into similar problems, by the way, if we employ the formulae (3) approach.

V. The Sign of the Coefficient and the Impact of Income Effects on Measures of Closeness

So far, we have discussed the absolute value of the coefficient of cross elasticity as a measure for the closeness of substitutes and we have taken for granted that the sign of the coefficient is positive for substitutes and negative for complements. Some textbook authors, e.g. Brennan [4, p. 90] and Mansfield [11, p. 93] warn that this need not be so if we have the Hicksian definitions of substitutes and complements in mind. Hicks himself [8, pp. 48-49, cases (2) and (3)] points out that the income effect of a fall in the price of a commodity X can be responsible for cases where a commodity

Y which, by his definition, is "mildly complementary with X" appears to be a substitute in terms of the total effect, provided Y is an inferior good and the income effect on Y is larger than the substitution effect. Alternatively, a commodity Y which is "mildly substitutable for X" can appear to be a complement, i.e. the demand for Y increases when the price of X falls, provided Y is a normal good and the income effect on Y outweighs the substitution effect. Or the effects going in opposite directions might just cancel out.¹² Hicks [8, pp. 49-50] "feels that a good deal of mild substitutability must be present which is prevented from showing itself by being offset by income effects."

Thus, due to income effects of price changes, we might find "false substitutes" and "false complements". To the extent that non-negligible income effects occur, how do they affect the use of cross elasticities for the purposes discussed? We can only measure the total effects on demand of a price change. What effect do we intend to measure for the purpose of measuring closeness or isolation? If not the total effects, what difference does it make for the interpretation of the various types of coefficients of cross elasticity? Take first the case

12. Note that the argument of this section is phrased in terms of an individual consumer's demand. But the conclusions also apply for market relationships. We can aggregate the individual demand curves even if X and Y may be regarded as complements by some consumers and as substitutes by others.

of a false substitute, Hicks' case (2). Commodity Y is a complement to commodity X by the Hicksian definition and an inferior good. When the price of X falls, the positive effect on the demand for Y is reduced by an income effect in the opposite direction. The negative income effect means that consumers replace Y by a third good, a superior good, when their income rises. However, consumers do not replace Y by its complement X ! If the income effect is strong enough to result in a negative total effect, the positive sign of any coefficient of cross elasticity between Y and X would therefore be completely misleading. It seems, however, that for practical purposes we can disregard the case of false substitutes (which might also be called "Giffen complements"), because they are supposed to be as rare as Giffen substitutes.

The case of false complements, Hicks' case (3), is believed to be more important. If Y is a substitute for X and the price of X is reduced, the substitution effect is positive for X and negative for Y because buyers switch to the good which has become relatively cheaper. The income effects are positive for both goods, provided both are normal goods. The total effect on the demand for X would be larger than the pure substitution effect; the total effect on the demand for Y could be negative but smaller than the pure substitution effect

(case of underrated substitute), or the total effect on the demand for Y could be positive (case of false complement). The damage that the occurrence of income effects could do to the various concepts discussed depends on the interpretation we give to our formulae. As long as we are only concerned with changes in the demand for one commodity at a time, the total effects seems to be what we want to measure. However, as soon as we think in terms of sales "gained from" or "lost to" other markets, income effects will create difficulties in measuring the sought effects. I already pointed this out for Krelle's coefficient of mobility, our formulae (4), in the previous section.

As regards the cross elasticity formulae (1) - (3), we seem to be only concerned with the effects on the demand for one commodity at a time, viz. the commodity with an unchanged price, when we determine a coefficient. This interpretation seems to be appropriate when we employ formulae (1) and (2) as criteria for isolation. One should think that the affected sellers are exclusively concerned with the per-cent changes in their own sales which result from income and substitution effects combined. If the total effects of a price change on sellers with unchanged prices are negligible there is neither a competitive nor a complementary relationship between the

price-changing seller and the others. If the coefficients are non-negligible and positive, the price changing seller is not isolated from the other sellers. I would even say that higher positive coefficients can be interpreted as reflecting a higher degree of interdependence of sellers, no matter what the pure substitution effects would be if we could compensate for the income effects. It is implied in this interpretation that the affected sellers' reactions depend exclusively on the changes in their own sales and that they do not consider the effects on the price-changing seller's sales. If this implication is not intended, the measure of isolation is also distorted by income effects.

When we employ formulae (1) and (3) as measures for the relative closeness of substitutes, it seems to me that a comparison of sales "gained from" or "lost to" other markets is always intended although the definitions of the coefficients only contain quantities of one commodity, the commodity with the unchanged price. When we speak of "closeness" of substitutes we actually visualize customers switching from one commodity to another; the closer a commodity is to its substitute the easier consumers will switch; thus closeness is measured by the quantities of sales moving between the markets in response to a given change in the price of one commodity. In

formulae (4) the "gained from" or "lost to" interpretation is made explicit in the definition of Krelle's coefficients of mobility. For any of the two alternative formulae (1) and (3), the pure substitution effects would be the best measure of closeness. As far as a problem of units exists, conversion factors would be some kind of weighted-average rate of substitution over a certain arc of the consumers' indifference curves. To the extent that income effects occur, the sales lost and gained differ from each other for each individual pair of substitutes. Comparisons of relative closeness would ordinarily be distorted as well. On the other hand, it seems that for most purposes for which the affected industries would want to know the relative closeness of substitutes the total effects on their sales are still an appropriate measure. Hicks concedes [8, p. 50] that for market relationships his distinction between substitutes and complements is less pure anyhow, because it is possible that two commodities may be regarded as complements by some consumers and as substitutes by others. It should also be realized that for very close substitutes, Hicks' case (4), the income effect is likely to be relatively unimportant. In cases of extreme substitutability we must expect corner solutions, i.e. some consumers completely stop buying the commodity which becomes more

expensive than a close substitute.

There is another question regarding the sign of the coefficients which is closely related to the case of "false complements". If we look at the relationship between a single commodity X and all other goods lumped together (these other goods have given prices and their quantities are measured in terms of a common yardstick commodity "money"), X and the commodity "money" then must be substitutes;¹³ X and money are likely to be mildly substitutable; in general money will be a normal good in the sense that the expenditure on all other goods taken together increases when income rises. Thus we can easily have the case indicated above that the income effect on money outweighs the substitution effect, i.e. the expenditure on all other goods increases when the price of X falls. The coefficient of cross elasticity in this case would be negative, designating money as a "false complement". This case, of course, is a familiar one. That the expenditure on other goods rises when the price of X falls implies that the expenditure on X falls. Thus the case of money as a "false complement" is the case of inelastic demand for good X .

This finding raises the question: Can we have positive coefficients of cross elasticity (substitutes for X) at all if the demand for X is inelastic? Triffin

13. Cf. Hicks [8, p. 46] or [9, p. 129].

[17, p. 132] seems to have been puzzled by this problem. The answer is in Hicks [8, pp. 46-47 and 51];¹⁴ What we find for the commodity "money" need not be true for each commodity in the group of other goods taken separately. When the price of good X with inelastic demand is lowered, consumers can still rearrange their expenditure in such a way that the total effects on some other commodities lead to a decrease in demand (substitutes in terms of total effects), while others show an increase in demand (complements in terms of total effects). For the individual goods then applies what I said about different cases of income effects in the preceding paragraphs.

VI. The Use of Cross Elasticities in Antitrust Cases

According to Singer [15, p. 56] the use of the concept of cross elasticity of demand by the Supreme Court in the Cellophane Case (1956) marks "a high point in the use of theoretical economic concepts in judicial antitrust opinions". Other authors, e.g. Stocking [16, p. 570], are more sceptical. I want to examine whether the concepts which were employed by the courts were appropriate for the purpose intended. First the purpose:

14. A first line of defense could have been to argue that sellers facing downward sloping demand curves would be predicted to operate along the elastic portion of their demand curves. However, this argument would not answer the question nor could it cover the cases of competitive industries and price control resulting in demand to be inelastic at the given prices.

The concept of cross price elasticity was (and presumably is) one of several tests employed to define the "relevant market" in cases which either involve monopolization charges under Section 2 of the Sherman Act or merger cases covered by Section 7 of the Clayton Act (as amended in 1950). Under a monopolization charge, the defendant is interested in having the relevant market defined as widely as possible in order to reduce the significance of his own share in such a market. In merger cases, the defendants benefit when the definition is either so broad that no substantial market share is affected by the merger, or so narrow that the merging firms are found to operate in separate markets.

Whereas in important earlier monopoly cases, e.g. the Alcoa Case (1945), markets were narrowly defined, in the Cellophane Case the courts tried to face the facts of monopolistic competition, i.e. the availability of substitutes for the product of a seller who is charged with monopolization. Supreme Court Justice Reed described the problem in the following way:

Determination of the competitive market for commodities depends on how different from one another are the offered commodities in character or use, how far buyers will go to substitute one commodity for another....where there are market alternatives that buyers may readily use for their purposes, illegal monopoly does not exist merely because the product said to be monopolized differs from others. If it were not so, only physically identical products would be a part of the market.¹⁵

15. 351 U.S. 377 (1956), pp. 393-394.

The question was where to draw the line between substitutes to be included in and those to be excluded from the relevant market for the product under consideration. As courts do, the Supreme Court developed a doctrine: The relevant market "is composed of products that have reasonable interchangeability for the purposes for which they are produced -- price, use and qualities considered".¹⁶ In a multifarious approach, the court found that other flexible wrapping materials were reasonably interchangeable with cellophane. Cellophane made by DuPont only had a share of about 20 percent in the market of all flexible packaging materials, a figure which was insufficient to establish the requisite market power for the asserted monopolization charge.

It is important to note that the test of "reasonable interchangeability" comprises a detailed examination of various price and quality characteristics of potential substitutes, of their interchangeability in important end-uses, as well as of sellers' policies in promoting their products in competition with each other.¹⁷ The courts draw the boundaries of the relevant market where they feel that it is no longer reasonable to speak of interchangeable products. Presumably they are looking

16. Ibidem, p. 404.

17. In this respect there are strong similarities between the Antitrust consideration of interchangeability and the U.N. study of steel substitutes [18] referred to above.

for a significant "gap" in the chain of substitutes. In the Cellophane Case, there seemed to be no doubt that a significant gap existed between flexible wrapping materials and other packaging materials. (To my knowledge, non-packaging materials were not even considered.) The issue at stake was whether there were any significant gaps between cellophane and other flexible wrapping materials.

Now the use of the concept of cross elasticity in pursuing the described purpose: By my reading, the courts use the terms "interchangeability" and "cross-elasticity of demand" largely as synonyma. It seems understandable that judges would hesitate to rely on a term such as "cross elasticity" to carry the message of their new doctrine. In one instance, Supreme Court Justice Reed specifically referred to the concept of cross price elasticity of demand. The following quote seems important, and it can serve as a starting point for an application of the ideas developed earlier:

An element for consideration as to cross-elasticity of demand between products is the responsiveness of sales of one product to price changes of the other. If a slight decrease in the price of cellophane causes a considerable number of customers of flexible wrappings to switch to cellophane, it would be an indication that a high cross-elasticity of demand exists between them; that the products compete in the same market.¹⁸

18. 351 U.S. 377 (1956), p. 400.

In effect, Justice Reed accepted the concept of cross price elasticity as a measure amongst others for the interchangeability of substitutes. I have not yet found any indication in the Supreme Court's decision nor the literature on the Cellophane case (or any other case for that matter) that the courts actually evaluated quantitative estimates of cross price elasticity. Let us assume for the sake of an argument that coefficients of cross price elasticity could have been estimated in a procedure which takes care of the ceteris paribus conditions and that the courts would have admitted such evidence. Would the test suggested by Justice Reed have led to an unambiguous result? In a charge under Section 2 of the Sherman Act, it seems reasonable to consider the potential own price policies of the alleged monopolist rather than price changes by others. That is what Justice Reed did. If he was referring to cross price elasticities defined in the usual way, he would have had our formulae (2) approach in mind. The wording of his statement, however, is consistent with the formulae (3) approach, because he spoke of customers switching to cellophane rather than of relative changes in the sales of substitute materials. Let us compare the implications of alternative approaches for the definition of a relevant market.

I showed in Section II why formulae (2) are a misleading approach to measuring the relative closeness of substitutes for a product. In the present context consider the following example: We find that a substitute called "glass film" increases its sales by 20 per cent when the price of cellophane is raised by one per cent and that the sales of another substitute, "vinyle", increase by only 5 per cent. If the Court had ruled that products to be included in the same market with cellophane must have a coefficient of not lower than 10, vinyle would have been excluded. And yet, 5 per cent of the vinyle market might be a much larger quantity than 20 per cent of relatively small glass film sales; the Court thus might have excluded a product which in terms of buyers switching is more important as a substitute for cellophane than products included in the relevant market. If before the price increase the cellophane sales were small relative to the total vinyle sales or if there were a large number of good substitutes, it could happen that because of a small price increase cellophane sales are severely reduced and still the formulae (2) coefficients could all turn out to be of negligible size. In such cases, would we want to conclude that cellophane has no close substitutes, therefore the product has a market for itself and the producer is a monopolist if he is the only producer of cellophane? This is the same problem which

Chamberlin [6, p. 265] faced when he found that low coefficients are common for all cases of isolated selling. He proposed to use the own elasticity of demand as an additional criterion. Brown and Wiseman [5, pp. 83-84] explicitly employ the misleading formulae (2) approach for the purpose of defining close substitutes. I repeat that the courts did not do so explicitly; and in any event, they relied on a much broader set of tests.

We saw in Section IV above that formulae (3) and (4) might be acceptable alternatives to the formulae (2) approach if we intend to determine the relative closeness of substitutes. However, for the purpose of defining a relevant market, formulae (3) and (4) would be ambiguous as well. In continuation of the above example let us assume that formulae (4) would lead to the result that cellophane loses one percent of its market to glass film when the price of cellophane, *ceteris paribus*, is increased by one percent and that cellophane at the same time loses 15 percent of its sales to vinyle producers. If the Court was looking for close substitutes and if it had used a coefficient of, say, 10 as a cut-off point, vinyle would have been included in the relevant market and glass film would not have been. This result seems unsatisfactory because glass film, which on the basis of formulae (2) increases its

sales by 20 per cent, should have been counted. If glass film had been grouped together with some other small close substitutes it might have been counted. The result thus depends on how narrowly we define the other products. Regarding vinyle, the result is also unsatisfactory because the relevant market would be "diluted" or "inflated" by the large volume of vinyle sales which in our example overlap only to a small degree with cellophane sales. If 5 per cent of the vinyle sales equal 15 per cent of the cellophane sales before the price increase, cellophane has a "market share" of only 25 per cent when the total sales of the two goods combined are declared to constitute the relevant market. One would have to employ additional tests, e.g. by consideration of end-uses, in order to separate the overlapping parts of the markets from the non-competing sales.¹⁹

The upshot is that the concept of cross price elasticity would be useless for the intended purpose even if problems of measuring the coefficients could be overcome. Certainly, one might argue that the formulae (4) test might provide valuable insights for the evaluation of an alleged monopoly position. It makes a differ-

19. Analogous problems arise with the formulae (3) approach and, incidentally, also with formulae (1). The latter approach was not discussed in this context because it puts the emphasis on other seller's price policies.

ence whether a cellophane producer contemplating a price increase must expect to lose 15 per cent of his sales to vinyl or only 1.5 per cent, and in merger cases policy makers might be the more concerned the higher are the coefficients between the sales of the merger candidates. However, this information by itself does not help in defining the relevant market in which the alleged monopolist's "market power" could be established. The reason that formulae (4) seem to convey information relevant for the monopoly problem is that the coefficients are defined as sectoral own price elasticities and that they add up to the monopolist's own elasticities. Chamberlin [6, p. 258] and others have argued that the own price elasticity cannot be used to measure the degree of monopoly power. The courts in monopoly and merger cases are not interested in the total percentage of sales a seller loses when he increases his price by one per-cent. They do not consider how many customers buy "something else". They are rather interested in the availability of well recognizable substitutes which buyers may use for the same purposes.²⁰ This approach, by necessity, includes too little and too much when the relevant market is defined.

20. Several formulations in the Supreme Court Decision and particularly in the minority opinion in the Cellophane Case suggest that the courts were guided by a dynamic concept of oligopolistic competition and that they were looking for markets where producers react to each others moves and where buyers could play off individual sellers against each other.

For the other potential uses of cross elasticities there was some doubt whether they should be employed as "pure demand" relationships or whether considerations of cost and supply should enter the argument; in the context of defining a relevant market, supply considerations seems to be indispensable. M.A. Adelman [1, pp. 688-690] has argued this point very strongly in connection with the Bethlehem-Youngstown Case (1958), the first merger case under the revised Section 7 of the Clayton Act. We could, for instance, find that the cross price elasticity of demand between re-inforcing bars and track spikes is negligible by any definition; the characteristics of the finished products are specific enough that a test of interchangeability in end-uses would classify them as separate lines of commerce. However, the technology of production seems to be such that a steel company with a basic bar mill can produce not only hot rolled bars, but also track spikes by adding a track spike machine at relatively low capital cost to its bar mill. Thus the potential cross price elasticity of demand for the basic output, hot rolled steel, can be substantial although the interchangeability of the finished product is negligible. Other examples would be boys' and girls' shoes or left-handed and right-handed golf clubs. In any instance, the potential competitors who could rapidly shift production towards the more

profitable output when relative prices change as well as the potential entry of new or imported substitutes must be considered when the boundaries of the relevant markets are determined for monopoly and merger cases.

REFERENCES

- [1] M.A. Adelman, "Economic Aspects of the Bethlehem Opinion", Virginia Law Review, Vol. 45 (1959), pp. 684-696.
- [2] Robert L. Bishop, "Elasticities, Cross-Elasticities, and Market Relationships", American Economic Review, Vol. 42, No. 5, (December 1952), pp. 779-803.
- [3] Kenneth E. Boulding, Economic Analysis, Volume 1: Microeconomics; fourth edition, New York: Harper & Row, 1966.
- [4] Michael J. Brennan, Theory of Economic Statics, Second Edition, Englewood Cliffs: Prentice Hall, 1970.
- [5] E.H. Phelps Brown and J. Wiseman, A Course in Applied Economics, 2nd ed., London: Pitman, 1964.
- [6] Edward H. Chamberlin, "Measuring the Degree of Monopoly and Competition", Monopoly and Competition and Their Regulation, Papers and Proceedings of a Conference held by the International Economics Association in September 1951, edited by E.H. Chamberlin; London: Macmillan, 1954, pp. 255-267.
- [7] Edward H. Chamberlin, "Supplementary Note on Cross-Elasticity" in his Towards a More General Theory of Value, New York: Oxford University Press, 1957, pp. 84-91.
- [8] J.R. Hicks, Value and Capital, 2nd ed., Oxford: Oxford University Press, 1946.
- [9] J.R. Hicks, A Revision of Demand Theory, Oxford: Oxford University Press, 1956.
- [10] Wilhelm Krelle, Preistheorie, Tübingen: Mohr, 1961.
- [11] Edwin Mansfield, Microeconomics, New York: Norton, 1970.
- [12] Ralph W. Pfouts and C.E. Ferguson, "Market Classification Systems in Theory and Policy", Southern Economic Journal, Vol. 26 (October 1959), pp. 111-118.

- [13] Paul A. Samuelson, Foundations of Economic Analysis. Paperback edition; New York: Antheneum, 1965.
- [14] Frederic M. Scherer, Industrial Market Structure and Economic Performance, Chicago: Rand McNally & Co., 1970.
- [15] Eugene M. Singer, Antitrust Economics: Selected Legal Cases and Economic Models, Englewood Cliffs: Prentice Hall, 1968.
- [16] George W. Stocking, "Economic Change and the Sherman Act: Some Reflections on 'Workable Competition'," Virginia Law Review, Vol. 44, No. 4 (May 1958), pp. 537-582.
- [17] Robert Triffin, Monopolistic Competition and General Equilibrium Theory, Cambridge, Mass.: Harvard University Press, 1962.
- [18] United Nations, Aspects of Competition Between Steel and Other Materials, New York: United Nations, 1966.
- [19] United States v. E.I. Du Pont De Nemours & Co., 351 U.S. 377 (1956).